Sepulveda Gil Susana Judith 27/02/20

27/02/20

# Media

## ▶ Media Aritmetica

Es la suma de todos los datos de la variable dividida por él número total de observaciones.

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i = \frac{1}{n}(X_1 + X_2 + \ldots + X_n) \text{ donde}$$

$X_i$: es el valor particular que registra el individuo i.

Ejemplo: sea X el número de días que tarda la producción de un pedido en la empresa A.

| Cliente | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| X | 10 | 1 | 7 | 16 | 0 | 2 | 3 | 5 | 10 | 6 | 0 | 8 |

Donde n = 12

$$\overline{X} = \frac{1}{12}\sum_{i=1}^{n} x_i = \frac{1}{12}(10+1+7+16+0+2+3+5+10+6+0+8) = 5.67$$

Entonces el promedio de los días que tarda en despachar el pedido es 5.67 días.

## ▶ Media ponderada

Es la suma de los valores $X_i$ de una Variable, multiplicada por sus ponderaciones $w_i$ (grado de importancia o peso de cada valor) y dividida por la suma de todas las ponderaciones.

Para una serie de datos $X = \{X_1, X_2, \ldots, X_n\}$ a la que corresponden los pesos $W = \{w_1, w_2, \ldots, w_n\}$ la media media ponderada se calcula como:

$$X_p = \frac{\sum_{i=1}^{n} w_i \cdot x_i}{\sum_{i=1}^{n} w_i} \quad o \quad X_p = \frac{w_1 \cdot x_1 + w_2 \cdot x_2 + w_3 \cdot x_3 + \ldots + w_n \cdot x_n}{w_1 + w_2 + w_3 + \ldots + w_n}$$

Con i = 1 ... n

Ejemplo: Según el caso anterior ¿cuál es el promedio ponderado de los días que tarda la empresa en entregar los pedidos?

| Cliente | 5 | 11 | 2 | 6 | 7 | 8 | 10 | 3 | 12 | 1 | 9 | 4 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $X_i$ | 0 | 0 | 1 | 2 | 3 | 5 | 6 | 7 | 8 | 10 | 10 | 16 |
| $W_i$ | 0.10 | 0.10 | 0.10 | 0.20 | 0.15 | 0.10 | 0.03 | 0.05 | 0.03 | 0.10 | 0.02 | 0.02 |

Se multiplican las ponderaciones por los valores de cada X:

| Cliente | $X_i$ | $W_i$ | $X_iW_i$ |
|---|---|---|---|
| 5 | 0 | 0.1 | 0 |
| 11 | 0 | 0.1 | 0 |
| 2 | 1 | 0.1 | 0.1 |
| 6 | 2 | 0.2 | 0.4 |
| 7 | 3 | 0.15 | 0.45 |
| 8 | 5 | 0.10 | 0.5 |
| 10 | 6 | 0.03 | 0.18 |
| 3 | 7 | 0.05 | 0.35 |
| 12 | 8 | 0.03 | 0.24 |
| 1 | 10 | 0.1 | 1 |
| 9 | 10 | 0.02 | 0.2 |
| 4 | 16 | 0.02 | 0.32 |
| | 5.67 | 1 | 3.74 |

$$X_p = \frac{0.1 \cdot 10 + 0.1 \cdot 1 + 7 \cdot 0.05 + \ldots + 0.08 \cdot 8}{0.1 + 0.1 + 0.05 + \ldots + 0.08} = \frac{3.74}{1} = 3.74$$

En conclución, dado que cada tiempo de entrega tiene diferentes ponderaciones, entonces el promedio de los días de entrega tarda un periodo de 3.74

## Mediana

Es el valor que se encuentra en la posición media después de ordenar los valores de la variable de mayor a menor o de menor a mayor. Si el número de datos es impar, la media es el valor que se encuentra en el centro de la serie ordenada. El centro de la serie ordenada es $L_{X_e} = \frac{n+1}{2}$, $L_{X_e}$, es la posición de donde se encuentra la mediana $X_e$ dentro

de una serie ordenada.
Si el número de datos es par, la mediana es el promedio entre los dos valores que se ubican en el centro de la serie.

$$L_1, L_2, \ldots, L_{\frac{n}{2}}, L_{\frac{n}{2}+1}, \ldots, L_n$$

Los dos posibles centros son: $L_{\frac{n}{2}}$ y $L_{\frac{n}{2}+1}$

Entonces la mediana para un conjunto de datos cuando se trata de una serie par es.

$$X_e = \frac{\left( X_{L_{\frac{n}{2}}} + X_{L_{\frac{n}{2}+1}} \right)}{2}$$

Ejemplo: sea X el número de días que tarda la producción de un pedido en la empresa A.

| Cliente | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| X | 10 | 1 | 7 | 16 | 0 | 2 | 3 | 5 | 10 | 6 | 0 | 8 |

Serie Ordenada:

| L | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Cliente | 5 | 11 | 2 | 6 | 7 | 8 | 10 | 3 | 12 | 1 | 9 | 4 |
| X | 0 | 0 | 1 | 2 | 3 | 5 | 6 | 7 | 10 | 10 | | 16 |

Esta serie es par, por lo tanto la mediana se encuentra entre los valores ubicados en las posiciones 6 y 7, es decir, entre los valores 5 y 6. La mediana es el promedio entre 5 y 6, es decir 5.5

## Moda

Es el valor de la observación que aparece con mayor frecuencia.

Ejemplo: Sea X el número de días que tarda la producción de un pedido de la empresa A.

| Cliente | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| X | 10 | 1 | 7 | 16 | 0 | 2 | 3 | 5 | 10 | 6 | 0 | 8 |

Los números que más se repiten son 0 y 10, en este caso, y según el concepto de moda, los dos números pueden ser la moda; pero no se puede determinar cuál de los dos números es la moda, dado que hay más de un número que se repite con la misma frecuencia. Se dice que esta serie o variable es multimodal.

## Rango

Es la diferencia entre el valor maximo y el valor minimo de una variable.

Ejemplo: Sea X el número de días que tarda la producción de un pedido en la empresa A.

| Cliente | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| X | 10 | 1 | 7 | 16 | 0 | 2 | 3 | 5 | 10 | 6 | 0 | 8 |

maximo: 16
minimo: 0
El rango es $16 - 0 = 16$

## Varianza

Es la media aritmetica de las diferencias al cuadrado de cada observación en relación a la media.

- ~~Variable~~ Poblacional

$$\sigma^2 = \frac{\sum_{i=1}^{n} (x_i - \mu)^2}{N}$$

UPAK

Donde

$X_i$ = es el valor particular que registra la observación i.
$\mu$ = es el valor de la media poblacional de la variable X
N = es el total de las observaciones de la población.

La varianza poblacional se utiliza cuando se observa toda la población, pero si en su lugar se toma una muestra de la población para análisis, entonces se calcula la varianza muestral $S^2$

- Varianza muestral

$$S^2 = \frac{\sum_{i=1}^{n} (X_i = \bar{X})^2}{n-1}$$

Donde

$X_i$ = es el valor particular que registra la observación i
$\bar{X}$ = es el valor de la media muestral de la variable X
n = es el total de las observaciones de la muestra.

Ejemplo: Suponiendo que la empresa A tomo una muestra de pedidos y encontro que el número de días que tarda la producción de un pedido es:

| Cliente | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| X | 10 | 1 | 7 | 16 | 0 | 2 | 3 | 5 | 10 | 6 | 0 | 8 |

Hallar la varianza muestra

| $X_i$ | $X_i - \bar{X}$ | $(X_i - \bar{X})^2$ |
|---|---|---|
| 10 | 4.33 | 18.78 |
| 1 | -4.67 | 21.78 |
| 7 | 1.33 | 1.78 |
| 16 | 10.33 | 106.78 |
| 0 | -5.67 | 32.11 |
| 2 | -3.67 | 13.44 |
| 3 | -2.67 | 7.11 |
| 5 | -0.67 | 0.44 |
| 10 | 4.33 | 18.78 |
| 6 | 0.33 | 0.11 |

| $x_i$ | $x_i - \bar{x}$ | $(x_i - \bar{x})^2$ |
|---|---|---|
| 0 | -5.67 | 32.11 |
| 8 | 2.33 | 5.44 |
| | | 258.67 |

$$S^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1} = \frac{258.67}{12-1} = 23.52$$

Este resultado no tiene interpretación inmediata dado que la varianza se explora en unidades al cuadrado. en el caso del ejemplo la varianza muestral es de 23.52 días al cuadrado.

## Desviación estándar

Es la raiz cuadrada de la varianza.

- Desviación poblacional

$$\sigma = \sqrt{\sigma^2}$$

- Desviación muestral

$$S = \sqrt{S^2}$$

Esta medida se define para dar una interpretación a la varianza.

Ejemplo: Sea X el número de días que tarda la producción de un pedido en la empresa A. Entonces, la desviación estándar de esta variable es:

$$S = \sqrt{S^2} = \sqrt{23.52} = 4.84$$

Según el resultado, en promedio la empresa tarda en entregar un pedido 5.67 días, pero con una desviación estandar de 4.84 lo que indica que el promedio de entrega se encuentra entre 0.82 y 10.52 días.

# Cuartil

Los cuartiles dividen el número de observaciones en cuatro partes iguales, donde:

- El primer cuartil tiene el 25% de los datos
- El segundo cuartil tiene el 50% de los datos y la ubicación de la media
- El tercer cuartil tiene el 75% de los datos.
- El cuarto cuartil tiene el 100% de los datos.

El primer cuartil es el porcentual 25, entonces:

$$L_p = \frac{(n+1)P}{100}$$

$$L_{25} = \frac{(12+1) \cdot 25}{100} = 3.25$$

Esta posición se encuentra entre el tercer y cuarto valor, entonces:
Se calcula la diferencia entre los valores del cuarto y terser puesto $(2-1=1)$.
Se saca el 0.25 de esta diferencia $(.25 \cdot 1 = 0.25)$ lo que indica que el primer cuartil es 1.25, mayor a 1 y menor a 2.

## Bibliografía:


Herramientas de estadística I / Liliana Adriana Mendoza Saboya / Editorial Universidad del Rosario / ISBN: 978-958-8378-10-7


# Sigma

nos permite escribir una suma con muchos términos en la forma compacta.

$$\sum_{k=1}^{n} a_k = a_1 + a_2 + a_3 + \dots + a_{n-1} + a_n$$

la letra griega $\Sigma$ (sigma mayúscula, corresponde a nuestra letra S). Significa "suma". El índice de la sumatoria K nos dice en donde empieza la Suma (mediante el número que está debajo del símbolo $\Sigma$) y en dondé termina (usando el número que esta arriba del simbolo $\Sigma$). Se puede usar cualquier letra para denotar el índice, pero las letras más usuales son i, j y k.

— El índice K termina en $K=n$

El símbolo de la sumatoria (letra griega sigma) → $\sum_{k=1}^{n} a_k$ — $a_k$ es una formula del k-esimo término.

El índice k empieza en $k=1$

De acuerdo con ello, podemos escribir

$$1^2+2^2+3^2+4^2+5^2+6^2+7^2+8^2+9^2+10^2+11^2 = \sum_{k=1}^{11} k^2$$

y

$$f(1)+f(2)+f(3)+\ldots+f(100) = \sum_{i=1}^{100} f(i).$$

La notación sigma usada en el lado derecho de estas ecuaciones es mucho más compacta que la suma del lado izquierdo.

Ejemplo:

$$\sum_{k=1}^{5} k \quad 1+2+3+4+5 \quad = 15$$

Muy Bien

Bibliografia:

Cálculo una variable / Thomas / Editorial Pearson / ISBN - 970-26-0643-8-90000

Falto año